

RECEIVED

AUG 23 2001

TECH CENTER 1600/2900

1632

H-21

8-28-01

P. 2

RAW SEQUENCE LISTING

DATE: 07/24/2001

PATENT APPLICATION: US/09/175,683B

TIME: 08:36:27

Input Set : A:\10275-134001.TXT

Output Set: N:\CRF3\07242001\I175683B.raw

ENTERED

4 <110> APPLICANT: Chen, Li How
 5 Meade, Harry M.
 7 <120> TITLE OF INVENTION: NOVEL MODIFIED NUCLEIC ACID SEQUENCES
 8 AND METHODS FOR INCREASING MRNA LEVELS AND PROTEIN
 9 EXPRESSION IN CELL SYSTEMS
 12 <130> FILE REFERENCE: 10275-134001
 14 <140> CURRENT APPLICATION NUMBER: US 09/175,683B
 15 <141> CURRENT FILING DATE: 1998-10-20
 17 <150> PRIOR APPLICATION NUMBER: US 60/085,649
 18 <151> PRIOR FILING DATE: 1998-05-15
 20 <150> PRIOR APPLICATION NUMBER: US 60/062,592
 21 <151> PRIOR FILING DATE: 1997-10-20
 23 <160> NUMBER OF SEQ ID NOS: 19
 25 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 27 <210> SEQ ID NO: 1
 28 <211> LENGTH: 1065
 29 <212> TYPE: DNA
 30 <213> ORGANISM: Artificial Sequence
 32 <220> FEATURE:
 33 <223> OTHER INFORMATION: altered MSP sequence; preferably, a bacterium,
 34 virus, or parasite
 36 <221> NAME/KEY: CDS
 37 <222> LOCATION: (1)...(1065)
 39 <400> SEQUENCE: 1
 40 gcc gtc act ccc tcc gtc atc gat aac atc ctg tcc aag atc gag aac 48
 41 Ala Val Thr Pro Ser Val Ile Asp Asn Ile Leu Ser Lys Ile Glu Asn
 42 1 5 10 15
 44 gag tac gag gtg ctg tac ctg aag ccg ctg gca ggg gtc tac cgg agc 96
 45 Glu Tyr Glu Val Leu Tyr Leu Lys Pro Leu Ala Gly Val Tyr Arg Ser
 46 20 25 30
 48 ctg aag aag cag ctg gag aac aac gtg atg acc ttc aac gtg aac gtg 144
 49 Leu Lys Lys Gln Leu Glu Asn Asn Val Met Thr Phe Asn Val Asn Val
 50 35 40 45
 52 aag gat atc ctg aac agc cgg ttc aac aag cgg gag aac ttc aag aac 192
 53 Lys Asp Ile Leu Asn Ser Arg Phe Asn Lys Arg Glu Asn Phe Lys Asn
 54 50 55 60
 56 gtg ctg gag agc gat ctg atc ccc tac aag gat ctg acc agc agc aac 240
 57 Val Leu Glu Ser Asp Leu Ile Pro Tyr Lys Asp Leu Thr Ser Ser Asn
 58 65 70 75 80
 60 tac gtg gtc aag gat ccc tac aag ttc ctg aac aag gag aag aga gat 288
 61 Tyr Val Val Lys Asp Pro Tyr Lys Phe Leu Asn Lys Glu Lys Arg Asp
 62 85 90 95
 64 aag ttc ctg agc agt tac aac tac atc aag gat agc att gat acc gat 336
 65 Lys Phe Leu Ser Ser Tyr Asn Tyr Ile Lys Asp Ser Ile Asp Thr Asp
 66 100 105 110
 68 atc aac ttc gcc aac gat gtc ctg gga tac tac aag atc ctg tcc gag 384
 69 Ile Asn Phe Ala Asn Asp Val Leu Gly Tyr Tyr Lys Ile Leu Ser Glu

RAW SEQUENCE LISTING

DATE: 07/24/2001

PATENT APPLICATION: US/09/175,683B

TIME: 08:36:27

Input Set : A:\10275-134001.TXT

Output Set: N:\CRF3\07242001\I175683B.raw

```

70          115          120          125
72 aag tac aag agc gat ctg gat tca atc aag aag tac atc aac gat aag      432
73 Lys Tyr Lys Ser Asp Leu Asp Ser Ile Lys Lys Tyr Ile Asn Asp Lys
74      130          135          140
76 cag gga gag aac gag aag tac ctg ccc ttc ctg aac aac atc gag acc      480
77 Gln Gly Glu Asn Glu Lys Tyr Leu Pro Phe Leu Asn Asn Ile Glu Thr
78 145          150          155          160
80 ctg tac aag acc gtc aac gat aag att gat ctg ttc gtg atc cac ctg      528
81 Leu Tyr Lys Thr Val Asn Asp Lys Ile Asp Leu Phe Val Ile His Leu
82          165          170          175
84 gag gcc aag gtc ctg aac tac aca tat gag aag agc aac gtg gag gtc      576
85 Glu Ala Lys Val Leu Asn Tyr Thr Tyr Glu Lys Ser Asn Val Glu Val
86          180          185          190
88 aag atc aag gag ctg aat tac ctg aag acc atc cag gat aag ctg gcc      624
89 Lys Ile Lys Glu Leu Asn Tyr Leu Lys Thr Ile Gln Asp Lys Leu Ala
90      195          200          205
92 gat ttc aag aag aac aac aac ttc gtc ggg atc gcc gat ctg agc acc      672
93 Asp Phe Lys Lys Asn Asn Asn Phe Val Gly Ile Ala Asp Leu Ser Thr
94      210          215          220
96 gat tac aac cac aac aac ctg ctg acc aag ttc ctg agc acc ggt atg      720
97 Asp Tyr Asn His Asn Asn Leu Leu Thr Lys Phe Leu Ser Thr Gly Met
98 225          230          235          240
100 gtc ttc gaa aac ctg gcc aag acc gtc ctg agc aac ctg ctg gat ggg      768
101 Val Phe Glu Asn Leu Ala Lys Thr Val Leu Ser Asn Leu Leu Asp Gly
102          245          250          255
104 aac ctg cag ggg atg ctg aac atc agc cag cac cag tgt gtg aag aag      816
105 Asn Leu Gln Gly Met Leu Asn Ile Ser Gln His Gln Cys Val Lys Lys
106          260          265          270
108 cag tgt ccc cag aac agc ggg tgt ttc aga cac ctg gat gag aga gag      864
109 Gln Cys Pro Gln Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu
110          275          280          285
112 gag tgt aag tgt ctg ctg aac tac aag cag gaa ggt gat aag tgt gtg      912
113 Glu Cys Lys Cys Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val
114      290          295          300
116 gaa aac ccc aat cct act tgt aac gag aac aat ggt gga tgt gat gcc      960
117 Glu Asn Pro Asn Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala
118 305          310          315          320
120 gat gcc aag tgt acc gag gag gat tca ggg agc aac ggg aag aag atc      1008
121 Asp Ala Lys Cys Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile
122          325          330          335
124 acc tgt gag tgt acc aag cct gat tct tat cca ctg ttc gat ggt atc      1056
125 Thr Cys Glu Cys Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile
126          340          345          350
128 ttc tgt agt      1065
129 Phe Cys Ser
130          355
133 <210> SEQ ID NO: 2
134 <211> LENGTH: 1088
135 <212> TYPE: DNA

```

RAW SEQUENCE LISTING

DATE: 07/24/2001

PATENT APPLICATION: US/09/175,683B

TIME: 08:36:27

Input Set : A:\10275-134001.TXT

Output Set: N:\CRF3\07242001\I175683B.raw

```

136 <213> ORGANISM: Plasmodium falciparum
138 <220> FEATURE:
139 <221> NAME/KEY: CDS
140 <222> LOCATION: (1)...(1083)
142 <400> SEQUENCE: 2
143 gca gta act cct tcc gta att gat aac ata ctt tct aaa att gaa aat      48
144 Ala Val Thr Pro Ser Val Ile Asp Asn Ile Leu Ser Lys Ile Glu Asn
145 1 5 10 15
147 gaa tat gag gtt tta tat tta aaa cct tta gca ggt gtt tat aga agt      96
148 Glu Tyr Glu Val Leu Tyr Leu Lys Pro Leu Ala Gly Val Tyr Arg Ser
149 20 25 30
151 tta aaa aaa caa tta gaa aat aac gtt atg aca ttt aat gtt aat gtt      144
152 Leu Lys Lys Gln Leu Glu Asn Asn Val Met Thr Phe Asn Val Asn Val
153 35 40 45
155 aag gat att tta aat tca cga ttt aat aaa cgt gaa aat ttc aaa aat      192
156 Lys Asp Ile Leu Asn Ser Arg Phe Asn Lys Arg Glu Asn Phe Lys Asn
157 50 55 60
159 gtt tta gaa tca gat tta att cca tat aaa gat tta aca tca agt aat      240
160 Val Leu Glu Ser Asp Leu Ile Pro Tyr Lys Asp Leu Thr Ser Ser Asn
161 65 70 75 80
163 tat gtt gtc aaa gat cca tat aaa ttt ctt aat aaa gaa aaa aga gat      288
164 Tyr Val Val Lys Asp Pro Tyr Lys Phe Leu Asn Lys Glu Lys Arg Asp
165 85 90 95
167 aaa ttc tta agc agt tat aat tat att aag gat tca ata gat acg gat      336
168 Lys Phe Leu Ser Ser Tyr Asn Tyr Ile Lys Asp Ser Ile Asp Thr Asp
169 100 105 110
171 ata aat ttt gca aat gat gtt ctt gga tat tat aaa ata tta tcc gaa      384
172 Ile Asn Phe Ala Asn Asp Val Leu Gly Tyr Tyr Lys Ile Leu Ser Glu
173 115 120 125
175 aaa tat aaa tca gat tta gat tca att aaa aaa tat atc aac gac aaa      432
176 Lys Tyr Lys Ser Asp Leu Asp Ser Ile Lys Lys Tyr Ile Asn Asp Lys
177 130 135 140
179 caa ggt gaa aat gag aaa tac ctt ccc ttt tta aac aat att gag acc      480
180 Gln Gly Glu Asn Glu Lys Tyr Leu Pro Phe Leu Asn Asn Ile Glu Thr
181 145 150 155 160
183 tta tat aaa aca gtt aat gat aaa att gat tta ttt gta att cat tta      528
184 Leu Tyr Lys Thr Val Asn Asp Lys Ile Asp Leu Phe Val Ile His Leu
185 165 170 175
187 gaa gca aaa gtt cta aat tat aca tat gag aaa tca aac gta gaa gtt      576
188 Glu Ala Lys Val Leu Asn Tyr Thr Tyr Glu Lys Ser Asn Val Glu Val
189 180 185 190
191 aaa ata aaa gaa ctt aat tac tta aaa aca att caa gac aaa ttg gca      624
192 Lys Ile Lys Glu Leu Asn Tyr Leu Lys Thr Ile Gln Asp Lys Leu Ala
193 195 200 205
195 gat ttt aaa aaa aat aac aat ttc gtt gga att gct gat tta tca aca      672
196 Asp Phe Lys Lys Asn Asn Asn Phe Val Gly Ile Ala Asp Leu Ser Thr
197 210 215 220
199 gat tat aac cat aat aac tta ttg aca aag ttc ctt agt aca ggt atg      720
200 Asp Tyr Asn His Asn Asn Leu Leu Thr Lys Phe Leu Ser Thr Gly Met

```

RAW SEQUENCE LISTING

DATE: 07/24/2001

PATENT APPLICATION: US/09/175,683B

TIME: 08:36:27

Input Set : A:\10275-134001.TXT

Output Set: N:\CRF3\07242001\I175683B.raw

```

201 225          230          235          240
203 gtt ttt gaa aat ctt gct aaa acc gtt tta tct aat tta ctt gat gga      768
204 Val Phe Glu Asn Leu Ala Lys Thr Val Leu Ser Asn Leu Leu Asp Gly
205          245          250          255
207 aac ttg caa ggt atg tta aac att tca caa cac caa tgc gta aaa aaa      816
208 Asn Leu Gln Gly Met Leu Asn Ile Ser Gln His Gln Cys Val Lys Lys
209          260          265          270
211 caa tgt cca caa aat tct gga tgt ttc aga cat tta gat gaa aga gaa      864
212 Gln Cys Pro Gln Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu
213          275          280          285
215 gaa tgt aaa tgt tta tta aat tac aaa caa gaa ggt gat aaa tgt gtt      912
216 Glu Cys Lys Cys Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val
217          290          295          300
219 gaa aat cca aat cct act tgt aac gaa aat aat ggt gga tgt gat gca      960
220 Glu Asn Pro Asn Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala
221 305          310          315          320
223 gat gcc aaa tgt acc gaa gaa gat tca ggt agc aac gga aag aaa atc      1008
224 Asp Ala Lys Cys Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile
225          325          330          335
227 aca tgt gaa tgt act aaa cct gat tct tat cca ctt ttc gat ggt att      1056
228 Thr Cys Glu Cys Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile
229          340          345          350
231 ttc tgc agt cac cac cac cac cac taact      1088
232 Phe Cys Ser His His His His His His
233          355          360
236 <210> SEQ ID NO: 3
237 <211> LENGTH: 88
238 <212> TYPE: DNA
239 <213> ORGANISM: Artificial Sequence
241 <220> FEATURE:
242 <223> OTHER INFORMATION: synthetically generated oligonucleotide
244 <400> SEQUENCE: 3
245 tcgacgagag ccatgaaggt cctcaccctt gcctgtctgg tggctctggc cattgcaaga      60
246 gagcaggaag aactcaatgt agtcggta      88
248 <210> SEQ ID NO: 4
249 <211> LENGTH: 88
250 <212> TYPE: DNA
251 <213> ORGANISM: Artificial Sequence
253 <220> FEATURE:
254 <223> OTHER INFORMATION: synthetically generated oligonucleotide
256 <400> SEQUENCE: 4
257 gatctaccga ctacattgag ttcttctgc tctcttgcaa tggccagagc caccagacag      60
258 gcaaggatga ggaccttcac ggctctcg      88
260 <210> SEQ ID NO: 5
261 <211> LENGTH: 60
262 <212> TYPE: DNA
263 <213> ORGANISM: Artificial Sequence
265 <220> FEATURE:
266 <223> OTHER INFORMATION: oligonucleotide for PCR

```

RAW SEQUENCE LISTING

DATE: 07/24/2001

PATENT APPLICATION: US/09/175,683B

TIME: 08:36:27

Input Set : A:\10275-134001.TXT

Output Set: N:\CRF3\07242001\I175683B.raw

```

268 <400> SEQUENCE: 5
269 taactcgagc gaaccatgaa ggtcctcatc cttgcctgtc tgggtggctct ggccattgca      60
271 <210> SEQ ID NO: 6
272 <211> LENGTH: 48
273 <212> TYPE: DNA
274 <213> ORGANISM: Artificial Sequence
276 <220> FEATURE:
277 <223> OTHER INFORMATION: oligonucleotide for PCR
279 <400> SEQUENCE: 6
280 aattctcgag ttagtggtgg tgggtgggtg gactgcagaa ataccatc      48
282 <210> SEQ ID NO: 7
283 <211> LENGTH: 31
284 <212> TYPE: DNA
285 <213> ORGANISM: Artificial Sequence
287 <220> FEATURE:
288 <223> OTHER INFORMATION: oligonucleotide for PCR
290 <400> SEQUENCE: 7
291 aatagatctg cagtaactcc ttccgtaatt g      31
293 <210> SEQ ID NO: 8
294 <211> LENGTH: 1142
295 <212> TYPE: DNA
296 <213> ORGANISM: Artificial Sequence
298 <220> FEATURE:
299 <223> OTHER INFORMATION: altered MSP sequence; preferably, a bacterium,
300 virus, or parasite
302 <221> NAME/KEY: CDS
303 <222> LOCATION: (1)...(1142)
305 <400> SEQUENCE: 8
306 atg aag gtc ctc ata att gcc tgt ctg gtg gct ctg gcc att gca gcc      48
307 Met Lys Val Leu Ile Ile Ala Cys Leu Val Ala Leu Ala Ile Ala Ala
308 1 5 10 15
310 gtc act ccc tcc gtc atc gat aac atc ctg tcc aag atc gag aac gag      96
311 Val Thr Pro Ser Val Ile Asp Asn Ile Leu Ser Lys Ile Glu Asn Glu
312 20 25 30
314 tac gag gtg ctg tac ctg aag ccc ctg gca gga gtc tac agg agc ctg      144
315 Tyr Glu Val Leu Tyr Leu Lys Pro Leu Ala Gly Val Tyr Arg Ser Leu
316 35 40 45
318 aag aag cag ctg gag aac aac gtg atg acc ttc aac gtg aac gtg aag      192
319 Lys Lys Gln Leu Glu Asn Asn Val Met Thr Phe Asn Val Asn Val Lys
320 50 55 60
322 gat atc ctg aac agc agg ttc aac aag agg gag aac ttc aag aac gtg      240
323 Asp Ile Leu Asn Ser Arg Phe Asn Lys Arg Glu Asn Phe Lys Asn Val
324 65 70 75 80
326 ctg gag agc gat ctg atc ccc tac aag gat ctg acc agc agc aac tac      288
327 Leu Glu Ser Asp Leu Ile Pro Tyr Lys Asp Leu Thr Ser Ser Asn Tyr
328 85 90 95
330 gtg gtc aaa gat ccc tac aag ttc ctg aac aag gag aag aga gat aag      336
331 Val Val Lys Asp Pro Tyr Lys Phe Leu Asn Lys Glu Lys Arg Asp Lys
332 100 105 110

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/175,683B

DATE: 07/24/2001

TIME: 08:36:28

Input Set : A:\10275-134001.TXT

Output Set: N:\CRF3\07242001\I175683B.raw